*VL*fitter

a likelihood inference for the CKM matrix elements

L. Velasco-Sevilla [1]
University of Minnesota

N. Leonardo [2]
Massachusetts Institute of Technology

October 6, 2005

[1] liliana@physics.umn.edu
[2] leonardo@fnal.gov

1 Likelihood technique

A global fit to the CKM matrix elements may be performed using several methods [1, 2, 3, 4]. These treat in different ways the available information on the experimental and the theoretical uncertainties. We employ a Bayesian approach to construct a global inference function, from which probability intervals for the relevant parameters may be derived. The uncertainties are described in terms of pdf's which quantify the confidence on the values of the involved variables.

We first motivate the procedure for a single constraint; a more general description will follow. The oscillation frequencies of the neutral B meson systems can be related, within the SM, to the CKM parameters $\bar{\rho}$ and $\bar{\eta}$ through an equation of the type

$$(1-\bar{\rho})^2 + \bar{\eta}^2 = c \tag{1}$$

where c is a quantity formed of the experimentally measured Δm ($c \propto \Delta m_d, {\Delta m_s}^{-1}$), and of other theoretically determined parameters.

In the ideal case where c would be perfectly known, the constraint expressed by equation 1 would result in a curve in the $\bar{\rho}$-$\bar{\eta}$ plane, *i.e.* a circle of radius $\sqrt{c}$. The pdf describing our beliefs in the $\bar{\rho}$ and $\bar{\eta}$ values would be

$$f(\bar{\rho},\bar{\eta}|c) = \delta((1-\bar{\rho})^2 + \bar{\eta}^2 - c) \tag{2}$$

The points in the circumference would appear as likely. This would remain so in the absense of other experimental piece of information, or theoretical prejudice, which might exclude points outside a determined *physical region*, or in general lead to the assignement of different weights to the various points.

In a realistic case c is not known exactly, the available knowledge about its value being contained in a corresponding pdf, $f(c)$. This way, instead of a single circle, there is in reality an infinite collection of curves, each having a weight $f(c)$. The expected values for $\bar{\rho}$ and $\bar{\eta}$ are thus obtained from

$$f(\bar{\rho},\bar{\eta}) = \int f(\bar{\rho},\bar{\eta}|c)\ f(c)\ dc \tag{3}$$

Supposing a best experimental estimate for c would be given by $\hat{c}$, with uncertainty σ_c, and assuming a Gaussian distribution, the previous equation would take the form

$$\begin{aligned} f(\bar{\rho},\bar{\eta}) &= \int \delta((1-\bar{\rho})^2 + \bar{\eta}^2 - c)\frac{1}{\sqrt{2\pi}\sigma_c} e^{-\frac{1}{2}\left(\frac{c-\hat{c}}{\sigma_c}\right)^2}\ dc \\ &= \frac{1}{\sqrt{2\pi}\sigma_c} e^{-\frac{1}{2}\left(\frac{(1-\bar{\rho})^2-\bar{\eta}^2-\hat{c}}{\sigma_c}\right)^2} \end{aligned}$$

In a more general case, c may be formed from various input quantities $\{x_i\}$, denoted by $\mathbf{x}$, and generally described by a joint pdf $f(\mathbf{x})$; when the various x_i can be considered independent, the joint distribution simplifies to $f(\mathbf{x}) \sim \prod_i f(x_i)$. Denoting by $c(\mathbf{x})$ the dependency of c on the input quantities $\mathbf{x}$, $f(c)$ can generally be obtained as

$$f(c) = \int f(\mathbf{x})\delta(c - c(\mathbf{x}))\ d\mathbf{x}$$

We describe now more generally the procedure employed in this analysis. It involves the construction of a global inference, $\mathcal{L}$, relating $\bar{\rho}$, $\bar{\eta}$, the constraints $\mathbf{c} = \{c_j\}_{j=1}^M$, and the parameters $\mathbf{x} = \{x_i\}_{i=1}^N$. The various constraints $\mathbf{c}$, standing for Δm_d, $\Delta m_s/\Delta m_d$, ϵ_K, $|V_{ub}/V_{cd}|$, $sin(2\beta)$, may be expressed as

$$c_j = c_j(\bar{\rho}, \bar{\eta}; \mathbf{x}) \tag{4}$$

where the parameters $\mathbf{x}$ denote here all experimentally measured and theoretically calculated quantities on which $\mathbf{c}$ depend. The set of measured constraint values is represented by $\hat{\mathbf{c}} = \{\hat{c}_j\}_{j=1}^M$.

Making use of Bayes's theorem,

$$\begin{aligned}
\mathcal{L}(\bar{\rho}, \bar{\eta}, \mathbf{c}, \mathbf{x}|\hat{\mathbf{c}}) &\propto f(\hat{\mathbf{c}}|\bar{\rho}, \bar{\eta}, \mathbf{c}, \mathbf{x}) \cdot f(\mathbf{c}, \mathbf{x}, \bar{\rho}, \bar{\eta}) \\
&\propto f(\hat{\mathbf{c}}|\mathbf{c}) \cdot f(\mathbf{c}|\mathbf{x}, \bar{\rho}, \bar{\eta}) \cdot f(\mathbf{x}, \bar{\rho}, \bar{\eta}) \\
&\propto f(\hat{\mathbf{c}}|\mathbf{c}) \cdot \delta(\mathbf{c} - \mathbf{c}(\mathbf{x}, \bar{\rho}, \bar{\eta})) \cdot f(\mathbf{x}) \cdot f_0(\bar{\rho}, \bar{\eta})
\end{aligned} \tag{5}$$

Here $f_0(\bar{\rho}, \bar{\eta})$ is the *prior* distribution for $\bar{\rho}, \bar{\eta}$, which we take as uniform; $f(\mathbf{x})$ denotes similarly the prior joint pdf for parameters $\mathbf{x}$. In the derivation we have noted that c_j are uniquivocally determined, within the SM, from the values of $\bar{\rho}$, $\bar{\eta}$, and $\mathbf{x}$, and that $\hat{\mathbf{c}}$ depends on those parameters only through $\mathbf{c}$. Considering the independence of the various quantities, equation 5 becomes

$$\mathcal{L}(\bar{\rho}, \bar{\eta}, \mathbf{x}) \propto \prod_{j=1,M} f(\hat{c_j}|c_j(\bar{\rho}, \bar{\eta}, \mathbf{x})) \times \prod_{i=1,N} f_i(x_i) \tag{6}$$

where the constraints imposed by the δ-functions in the previous expression are assumed, and the prior, constant f_0 distribution was also ommited.

Equation 6 constitutes our sought-after global inference. Within the framework of Bayes statistics, and upon normalization, the left-hand side of the relation is the *posterior* pdf for the argument parameters. The probability distribution for any of the involved parameters can be achieved by integration over the remaining, sometimes also called *nuisance*, quantities.

In a Baeysian approach the various uncertainties are treated in a similar fashion, such that there is no conceptual distinction between those due to random fluctuations in the measurements, those about the parameters of the theory, or those associated to systematics of parameters known but with limited accuracy. Indeed, a systematic uncertainty on a parameter on which the measured constraints depend may be handled by adding the parameter to the collection $\mathbf{x}$.

We consider two models for describing the uncertainties. A Guassian model is chosen when the uncertainty is dominated by statistical effects, or there are many contributions to the systematics error, so that the central limit theorem applies. Otherwise, a uniform distribution is used for the uncertainty. When both a Gaussian and a flat uncertainty components are available for a parameter, the resulting pdf is obtained by convoluting the two distributions. I.e., for an observable parameter x of true value $\bar{x}$, with Gaussian and uniform uncertainty components, σ_g, σ_u, one has for the parameter and its pdf, $f(x)$,

$$\begin{aligned} x &= \bar{x} + x_g + x_u \\ f(x) &= \delta(x - \bar{x}) \otimes Gaus(x|\sigma_g) \otimes Unif(x|\sigma_f) \end{aligned}$$

Besides the constraints themselves, we classify the involved parameters into two classes: *(i) fitted*, for which we construct pdf's, and which are what we have been denoting by $\mathbf{x}$ (e.g. the top mass); and *(ii) fixed*, which are taken as constant (e.g. the W mass).

Joint pdf for $\bar{\rho}$-$\bar{\eta}$ and other *posterior* probabilities

The combined probability distribution for $\bar{\rho}$ and $\bar{\eta}$ are obtained by integrating Equation 6 over the (here *nuisance*) parameters $\mathbf{x}$,

$$\mathcal{L}(\bar{\rho},\bar{\eta}) \;\propto\; \int \prod_{j=1,M} f(\hat{c_j}|c_j(\bar{\rho},\bar{\eta},\{x_i\})) \times \prod_{i=1,N} f_i(x_i)\; dx_i \;\times f_0(\bar{\rho},\bar{\eta}) \tag{7}$$

The integration can be performed using Monte Carlo methods; then the normalization can be trivially performed, and all moments can also be easily computed. This expression shows explicitely that whereas *a priori* all values of $\bar{\rho}$ and $\bar{\eta}$ are equally likely by assumption, *i.e.* $f_0(\bar{\rho},\bar{\eta}) = const.$, *a posteriori* the probability clusters in a region of maximal likelihood.

The probability regions in the $\bar{\rho}$-$\bar{\eta}$ plane are constructed from the pdf obtained in equation 7. These are called *highest posterior density* regions, and are defined such that $\mathcal{L}(\bar{\rho},\bar{\eta})$ is higher everywhere inside the region than outside,

$$P_w := \{z = (\bar{\rho},\bar{\eta}) \;:\; \int_{P_w} \mathcal{L}(z)dz = w;\; \mathcal{L}(z') < min_{P_w}\mathcal{L}(z), \forall_{z' \notin P_w}\} \tag{8}$$

The single parameter pdf can also be obtained in the same fashion. For example, the $\bar{\rho}$ pdf is obtained as

$$\mathcal{L}(\bar{\rho}) \;\propto\; \int \mathcal{L}(\bar{\rho},\bar{\eta})d\bar{\eta} \tag{9}$$

from which its expected value can be calculated together with the corresponding highest posterior density intervals.

A similar procedure can be in principle used in order to obtain the pdf for other desired parameters. Technically, one may also use the probability function for transformed variables; *i.e.*, that for $\mathbf{u}(\mathbf{x})$ one has $f(\mathbf{u}) = f(\mathbf{x})|\partial\mathbf{x}/\partial\mathbf{u}|$, where the last factor denotes the Jacobian. This way, the pdf for a parameter x can be obtained from

$$\begin{aligned}\mathcal{L}(x) &\propto \int \mathcal{L}(x,\bar\eta)\ d\bar\eta \\ &\propto \int \mathcal{L}(\bar\rho,\bar\eta)\ |\frac{d\bar\rho}{dx}|\ d\bar\eta\end{aligned}$$

where $\mathcal{L}(\bar\rho,\bar\eta)$ has been computed in equation 7 above.

Besides the probability distribution in the $\bar\rho$-$\bar\eta$ plane, we are most interested in obtaining the *posterior* probability distribution for the Δm_s parameter.

2 Making use of the Δm_s amplitude information

The 95% C.L. exclusion limit, together with the sensitivity, provide a rather concise way of summarizing the results of the analysis. However, more information is contained in the full amplitude scan to the data. Therefore, in our fit to the CKM parameters we ought to use such more complete, continuous information about the degree of exclusion for Δm_s.

The measured values of the amplitude and its uncertainty, $\mathcal{A}$ and $\sigma_{\mathcal{A}}$, may be used to derive [**?**], in the Gaussian approximation, the log-likelihood function, $\Delta\ln\mathcal{L}^\infty(\Delta m_s)$, referenced to its value for an infinite oscillation frequency

$$\begin{aligned}\Delta\ln\mathcal{L}^\infty(\Delta m_s) &= ln\mathcal{L}(\infty) - ln\mathcal{L}(\Delta m_s) = \left(\frac{1}{2}-\mathcal{A}\right)\frac{1}{\sigma_{\mathcal{A}}^2} \\ \Delta\ln\mathcal{L}^\infty(\Delta m_s)_{mix} &= -\frac{1}{2}\frac{1}{\sigma_{\mathcal{A}}^2} \\ \Delta\ln\mathcal{L}^\infty(\Delta m_s)_{nomix} &= +\frac{1}{2}\frac{1}{\sigma_{\mathcal{A}}^2}\end{aligned}$$

The last two relations give the expected average log-likelihood value for the cases when Δm_s corresponds to the true (*mixing* case) or is far from (*no-mixing* case) oscillation frequency of the system, characterized respectively by unit and null expected amplitude values.

→**NB:** *Put the plots. ...*

The log-likelihood difference, according to the central limit theorem of likelihood theory, is χ^2-distributed, $\Delta ln\mathcal{L} = \frac{1}{2}\chi^2$. We translate therefore the amplitude scan into the likelihood ratio

$$R(\Delta m_s) = e^{-\Delta ln\mathcal{L}^\infty(\Delta m_s)} = \frac{\mathcal{L}(\Delta m_s)}{\mathcal{L}(\infty)} = e^{-\frac{\frac{1}{2}-\mathcal{A}(\Delta m_s)}{\sigma_{\mathcal{A}}^2(\Delta m_s)}} \tag{10}$$

through which the constraint for Δm_s is implemented in the fit. We re-iterate that the exponent in equation 10 corresponds to the χ^2, or log-likelihood, difference between the cases where an oscillation signal is present and absent, for which the true amplitude value is 1 and 0, respectively: $-\frac{1}{2}\left[\left(\frac{\mathcal{A}-1}{\sigma_{\mathcal{A}}}\right)^2-\left(\frac{\mathcal{A}}{\sigma_{\mathcal{A}}}\right)^2\right]$. Had the second term not been included, as it has been done in earlier studies, would result in Δm_s values with $\mathcal{A}>1$ being assigned lower probability with respect to those with $\mathcal{A}=1$, the undesirable. This way, hypotheses for Δm_s associated with larger $\mathcal{A}$-values in the scan contribute a larger weight in the fit.

→**NB:** *Can compare methods with toy, if time allows ...*

2.1 Extending the amplitude spectrum

The amplitude fits are performed for Δm_s values lower than a given threshold, beyond which the fit behavior may become unstable. On the other hand, in the frmaework of the CKM fit, it is in principle desirable to have R defined for all positive frequency values, which in turn demands for a continuation of the measured amplitude spectrum.

The extrapolation of the value of $\sigma_{\mathcal{A}}$ may be achieved, under the assumption of absense of a true oscillation singal in that region of the spectrum, through an analytical description of the significance curve

$$\sigma_{\mathcal{A}}^{-1} \propto W(\Delta m_s, \sigma_l, \sigma_p)$$

Here W in a known function describing the expected significance dependency on Δm_s thorugh the parameters σ_l and σ_p, which denote the B_s decay lenght and relative momentum uncertainties. For the combined amplitude scans from several mesurements, which is the case of the world average, these parameters may be adjusted using the measured part of the spectrum.

The extrapolation of A is not straighforward. The following possibilities may be considered:

- $\mathcal{A}=0$; which is the a priori expectation for no mixing.
- $\mathcal{A}=\frac{1}{2}$; this would set R to unit for the extrapolated Δm_s values.
- $\mathcal{A}\propto\sigma_{\mathcal{A}}$; assumption of constant $\frac{\mathcal{A}}{\sigma_{\mathcal{A}}}$ for Δm_s above sensitivity.

3 Constraints from neutral B meson mixing

Within the SM $\Delta m_{d,s}$ are very well aproximated by the relevant electroweak box diagrams. These are dominated by *top*-quark exchange. The result of the calculation obtained using the $\Delta B=2$ effective Hamiltonian yields for $|M_{12}^q|=\frac{1}{2}\Delta m_q$ $(q=d,s)$:

$$|M_{12}^q| = \frac{G_F^2}{12\pi^2}\eta_{B_q} m_{B_q}(B_{B_q}f_{B_q}^2)S(x_t)|V_{tq}^*V_{tb}|^2. \tag{11}$$

Here G_F is the Fermi constant; η_{B_q} is a QCD correction factor calculated in NLO; m_{B_q} and m_W are the B_q meson and W boson masses. The dominant uncertainties in equation (11) come from the evaluation of the hadronic quantities: f_{B_q}, the B meson decay constant, and B_{B_q}, which parameterizes the value of the hadronic matrix element. The Inami-Lim function is given by

$$S(x_t) \quad = \quad x_t \left[\frac{1}{4} + \frac{9}{4}\frac{1}{1-x_t} - \frac{3}{2}\frac{1}{(1-x_t)^2}\right] - \frac{3}{2}\left[\frac{x_t}{1-x_t}\right]^3 \ln x_t,$$

which describes the $|\Delta B| = 2$ transition amplitude in the absence of strong interaction, where the mass of the *top* quark enters via $x_t \equiv \frac{m_t^2}{M_W^2}$.

3.1 Δm_d

Expressing equation (11) above, for the B_d case, in terms of the Wolfenstein parameters, we obtain for Δm_d,

$$\Delta m_d = C_{\Delta m_d} A^2 \lambda^6 [(1-\bar{\rho})^2 + \bar{\eta}^2] m_{B_d} f_{B_d}^2 B_{B_d} \eta_{B_d} S(x_t), \tag{12}$$

where $C_{\Delta m_d} = \frac{G_F^2 M_W^2}{6\pi^2}$.

The parameters with dominant uncertainties in equation (12) are $f_{B_d}^2$, A, λ, which are varied parameters of the fit. A Gaussian constraint is implemented in the global likelihood

$$e^{-\frac{1}{2}\left(\frac{\Delta m_d - \Delta \hat{m}_d}{\sigma_{\Delta m_d}}\right)^2} \tag{13}$$

where Δm_d is provided by the r.h.s. of equation (12), and $\Delta \hat{m}_d$ and $\sigma_{\Delta m_d}$ denote the experimentally measured values.

3.2 Δm_s

The size of side $|V_{td}|/(\lambda|V_{cb}|)$ of the unitarity triangle can still be obtained from the ratio of $|M_{12}^d|$ and $|M_{12}^s|$:

$$\frac{\Delta m_d}{\Delta m_s} = \frac{|M_{12d}|}{|M_{12s}|} = \frac{m_{B_d} f_{B_d}^2 B_{B_d} \eta_{B_d}}{m_{B_s} f_{B_s}^2 B_{B_s} \eta_{B_s}} \frac{|V_{td}|^2}{|V_{ts}|^2}, \tag{14}$$

which is expected to be less dependent on the absolute values of f_B and B_B. Hence we can characterise it by

$$\xi = \frac{f_{B_s}\sqrt{B_{B_s}}}{f_{B_d}\sqrt{B_{B_d}}}, \tag{15}$$

the value of which is obtained from lattice QCD calculations.

The constraint we use from Δm_s is expressed, from equation (14), as

$$\Delta m_s = \Delta m_d \frac{m_{B_s}}{m_{B_d}} \xi^2 \frac{c^2}{\lambda^2} \frac{1}{(1-\bar{\rho})^2 + \bar{\eta}^2}, \tag{16}$$

where Δm_d is here taken as an experimental input.

The parameters with dominant uncertainties in equation (16) are $\xi,\ A,\ \lambda$, which are varied parameters of the fit. The constraint is implemented via the likelihood ratio, R, after accessing the amplitude point $(\mathcal{A}, \sigma_{\mathcal{A}})$ associated to the frequency value obtained by evaluating the r.h.s. of equation (16).

4 Other constraints and input

4.1 $|\frac{V_{ub}}{V_{cb}}|$

The CKM matrix elements $|V_{ub}|$ and $|V_{cb}|$ are measured in semileptonic B decays. A direct determination of the ratio is achieved via end point analysis in inclusive semileptonic B decays; the first is determined also from measurements of exclusive semileptonic branching ratios.

In terms of the re-scaled Wolfenstein parameters, the constraint $|V_{ub}|/|V_{cb}|$ is expressed as:

$$|V_{ub}|/|V_{cb}| = \frac{\lambda}{c}\sqrt{\bar{\rho}^2 + \bar{\eta}^2} \tag{17}$$

Both Gaussian and flat uncertainties are computed, and a corresponding convoluted pdf is employed in the implementation of the constraint.

4.2 $|\epsilon_K|$

The parameter ϵ_K expresses the measurement of indirect CP violation in the neutral K system.

In terms of the Wolfenstein parameters it is given by

$$\begin{aligned} |\epsilon_K| &= C_\epsilon B_K A^2 \lambda^6 \bar{\eta}[-\eta_1 x_c + A^2\lambda^4\left(1 - \bar{\rho} - \left(\bar{\rho}^2 + \bar{\eta}^2\right)\lambda^2\right)\eta_2 S(x_t) \\ &+ \eta_3 S(x_c, x_t)], \end{aligned} \tag{18}$$

where $C_\epsilon = \frac{G_F^2 f_K^2 m_K m_W^2}{6\sqrt{2}\pi^2 \Delta m_K}$. The short distance QCD corrections are codified in the the coefficients η_1, η_2 and η_3, and are functions of the charm and top quark masses and of the QCD scale parameter Λ_{QCD}; the η_i's have been calculated in the NLO. The Inami-Lim functions, which describe the $|\Delta S| = 2$ transition amplitude in the absence of strong interactions, are given by

$$\begin{aligned} S(x_t) &= x_t\left[\frac{1}{4} + \frac{9}{4}\frac{1}{1-x_t} - \frac{3}{2}\frac{1}{(1-x_t)^2}\right] - \frac{3}{2}\left[\frac{x_t}{1-x_t}\right]^3 \ln x_t, \\ S(x_c, x_t) &= -x_c \ln x_c + x_c\left[\frac{x_t^2 - 8x_t + 4}{4(1-x_t)^2}\ln x_t + \frac{3}{4}\frac{x_t}{x_t - 1}\right], \quad x_q \equiv \frac{m_q^2}{M_{\mathrm{W}}^2}. \end{aligned}$$

The parameters with dominant uncertainties are B_K, η_1, η_3, m_c and m_t.

A Gaussian constraint is implemented for $|\epsilon_K|$.

4.3 $\sin(2\beta)$

A direct determination of the angles of the unitarity triangle can be achieved via measurements of CP asymmetries in various B decays. The value of $\sin(2\beta)$ is meaasured in $B \to J/\psi K$ decays..

The UT angles can be expressed directly in terms of the re-scaled Wolfenstein parameters by:

$$\sin(2\beta) = \frac{2\bar{\eta}(1-\bar{\rho})}{\bar{\eta}^2 + (1-\bar{\rho})^2} \tag{19}$$

$$\sin(2\alpha) = \frac{2\bar{\eta}\left(\bar{\eta}^2 + \bar{\rho}(\bar{\rho}-1)\right)}{\left(\bar{\eta}^2 + (1-\bar{\rho})^2\right)\left(\bar{\eta}^2 + \bar{\rho}^2\right)} \tag{20}$$

$$\sin(2\gamma) = \frac{2\bar{\rho}\bar{\eta}}{\bar{\rho}^2 + \bar{\eta}^2} \tag{21}$$

A Gaussian constraint is implemented for $\sin(2\beta)$ alone. The expressions for $\sin(2\alpha)$ and $\sin(2\gamma)$ are also included, which may be used to obtain the corresponding posterior pdf's.

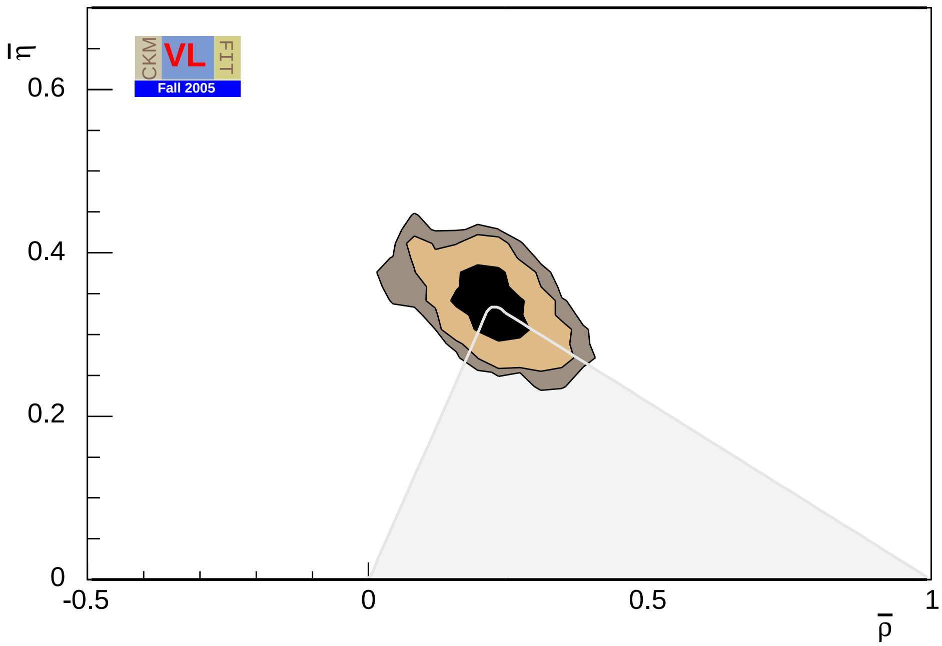
5 Posterior probability distributions

5.1 The unitarity triangle in the $\bar{\rho}$-$\bar{\eta}$ plane

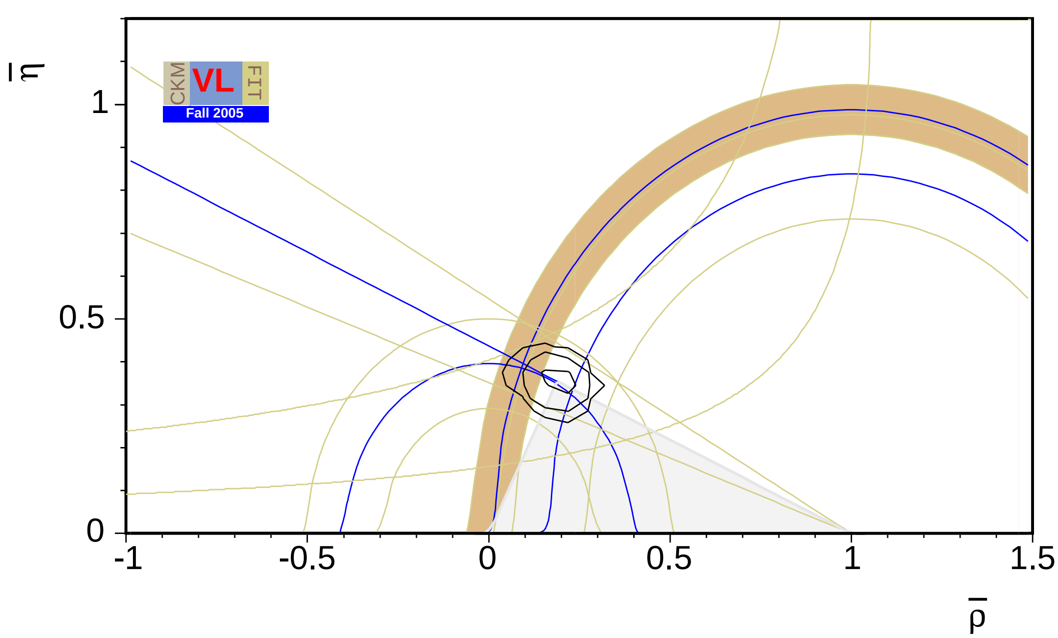
5.2 1-Dimensional posterior constraint distributions

5.3 Expected SM PDF for Δm_s

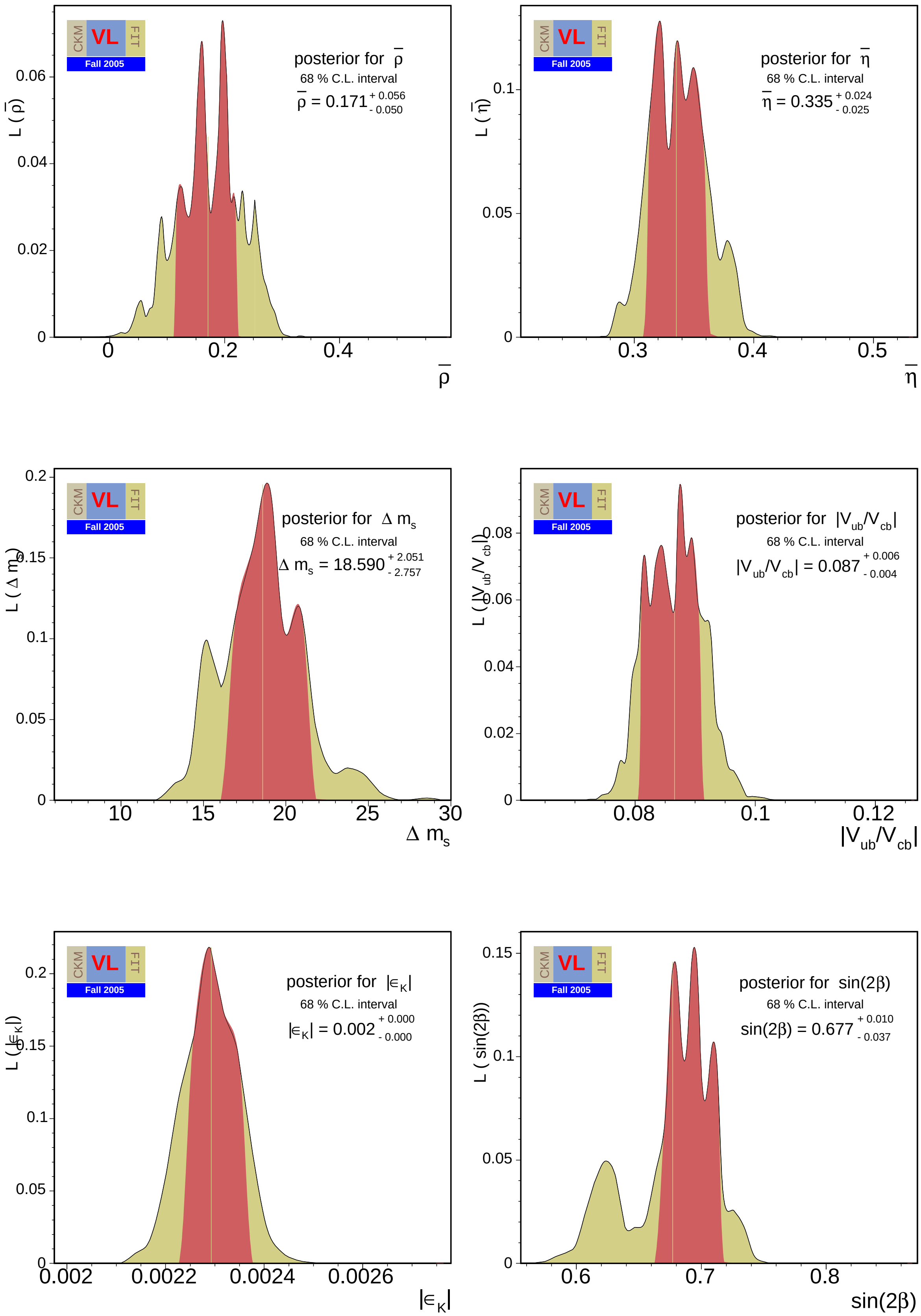
Here we show the probability distribution for Δm_s obtained from the SM fit. The fit is performed in two configurations, namely using and excluding the experimental information, provided by the amplitude scan, on the Δm_s parameter itself.



CKM
VL
FIT
Fall 2005
0.6
0.4
0.2
0
-0.5
0
0.5
1



CKM
VL
FIT
Fall 2005
1
0.5
0
-1
-0.5
0
0.5
1
1.5



CKM VL FIT
Fall 2005
posterior for ρ̄
68 % C.L. interval
ρ̄ = 0.171 +0.056 -0.050
L (ρ̄)
ρ̄
posterior for η̄
68 % C.L. interval
η̄ = 0.335 +0.024 -0.025
L (η̄)
η̄
posterior for Δ m_s
68 % C.L. interval
Δ m_s = 18.590 +2.051 -2.757
L (Δ m_s)
Δ m_s
posterior for |V_ub/V_cb|
68 % C.L. interval
|V_ub/V_cb| = 0.087 +0.006 -0.004
L (|V_ub/V_cb|)
|V_ub/V_cb|
posterior for |ε_K|
68 % C.L. interval
|ε_K| = 0.002 +0.000 -0.000
L (|ε_K|)
|ε_K|
posterior for sin(2β)
68 % C.L. interval
sin(2β) = 0.677 +0.010 -0.037
L (sin(2β))
sin(2β)

References

[1] M.Ciuchini, G.D'Agostini, E. Franco, V. Lubicz, G. Martinelli, F. Parodi, P. Roudeau, A. Stocchi, JHEP 0107 (2001) 013 [hep-ph/0012308].

[2] A. Höcker, H. Lacker, S. Laplace, F. Le Diberder, Eur. Phys. J. C**21** (2001) 225 [hep-ph/0104062].

[3] A. Ali, D. London, Eur. Phys. J. C**18** (2001) 665; S. Mele, Phys. Rev. D **59** (1999) 113011; D. Atwood, A. Soni, Phys. Lett. B **508** (2001) 17.

[4] Y. Grossman, Y. Nir, S. Plaszczynski, M.-H. Schune, Nucl. Phys. B **511** (1998) 69; S. Plaszczynski, M.-H. Schune [hep-ph/9911280].